



303 –LANGUAGE TEXT IDENTIFICATION

TEAM INFORMATION

Team Name: AWGN
 Results Email: [REDACTED]
 Examination Time Frame: 10/26 to 10/26/08

INSTRUCTIONS

Description: Examiners must develop and document a methodology used to identify the language presented in the documents in the **303_Language_Text_Identification_Challenge2008**.

Report the detailed explanation of your process (software or technique) used to determine the language and document information.

Total Weighted Points: 60 Total Points available per entry – Total 300 Points Available

1. **Answers** – Fill in the chart below with your findings. *As a Forensic Challenge, consider that your answers will have to have enough detail for the Findings and Methodology of your examination to satisfy questioning in a court of law.*
2. **Methodology** – Provide a meticulously detailed explanation of your process. Be sure to include a step action that our reviewers can follow to reproduce your work for authenticity including tools and techniques.

INTERNAL REVIEWER USE ONLY

Reviewer:

Points Awarded:

Date:

Review Period: to

Completed: ☐ Yes ☐ No ☐ Partial

Team AWGN 303

Page 1 of 6 11/18/2008



U.S. DEPARTMENT OF JUSTICE

U.S. DEPARTMENT OF JUSTICE

U.S. DEPARTMENT OF JUSTICE

U.S. DEPARTMENT OF JUSTICE

U.S. DEPARTMENT OF JUSTICE

U.S. DEPARTMENT OF JUSTICE

Report of Examination

Challenge Number: 303 - Language Text Identification

Examiner: Graham Eschbacher



Document1Tx.doc - Chinese	Google Translate
Document2Tx.doc - Russian	Google Translate
Document3Tx.doc - Korean	Google Translate
Document4Tx.doc - Arabic	Google Translate
Document5Tx.doc - Japanese	Google Translate

Examiner: Graham Eschbacher



Document Type: Chinese
Document Type: Korean
Document Type: Korean
Document Type: Arabic
Document Type: Japanese

Challenge Number: 303 - Language Text Identification**Tool Information**

Type	Name	Publisher
<input checked="" type="radio"/> Commercial <input type="radio"/> Open Source	Google Translate	Google
<input checked="" type="radio"/> Commercial <input type="radio"/> Open Source	Microsoft Office 2003 - Word	Microsoft
<input type="radio"/> Commercial <input type="radio"/> Open Source		
<input type="radio"/> Commercial <input type="radio"/> Open Source		
<input type="radio"/> Commercial <input type="radio"/> Open Source		

Notes

We used Google Translate (http://translate.google.com/translate_t?hl=en#) to identify the language in these text files. We simply opened the document in Microsoft Word (specifically Office Word 2003), copied and pasted the contents into the translator, and guessed the language until it was right.

